

存储计算融合为大数据添翼

Jeff Wu 吴建

海量增长的大数据需要被大量分析



10x↑



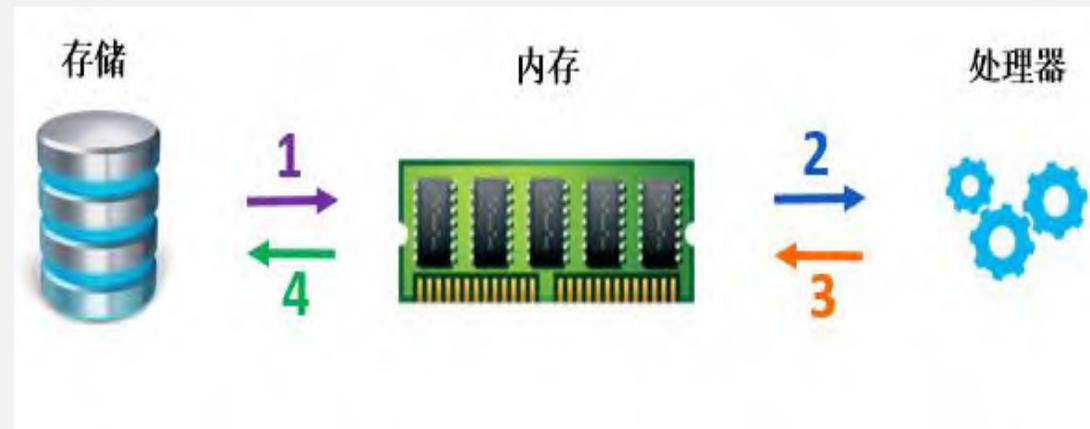
68%↑

存在问题:

- 17x 大数据量需要被分析，数据中心基础设施如何满足需求；
- IT管理人员面临着管理和扩展其基础设施以满足未来未知数据需求的艰巨任务。

数据中心基础设施的局限性

- 传统的“存储-内存-处理器”的服务器显示出若干局限性。
 - 在任何数量的数据可以被处理之前，它首先必须从存储转移到内存中；
 - 处理器接管加载，计算和操作数据的工作；
 - 然后将数据写回到内存，并最终传输回存储。



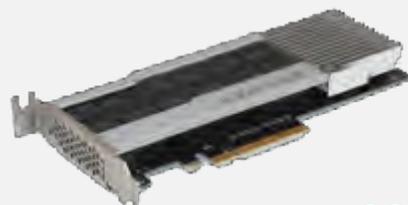
- 无法根据每个服务器需要处理的数据的速度和容量（每个横向扩展服务器上数百TB级别的数据）进行扩展以满足目前的需求，特别是当工作负载转移到实时响应级别。

当今闪存技术的问题



闪存或混合阵列

- 成本高昂
- 额外开销及冗余
- 实时分析时延过长

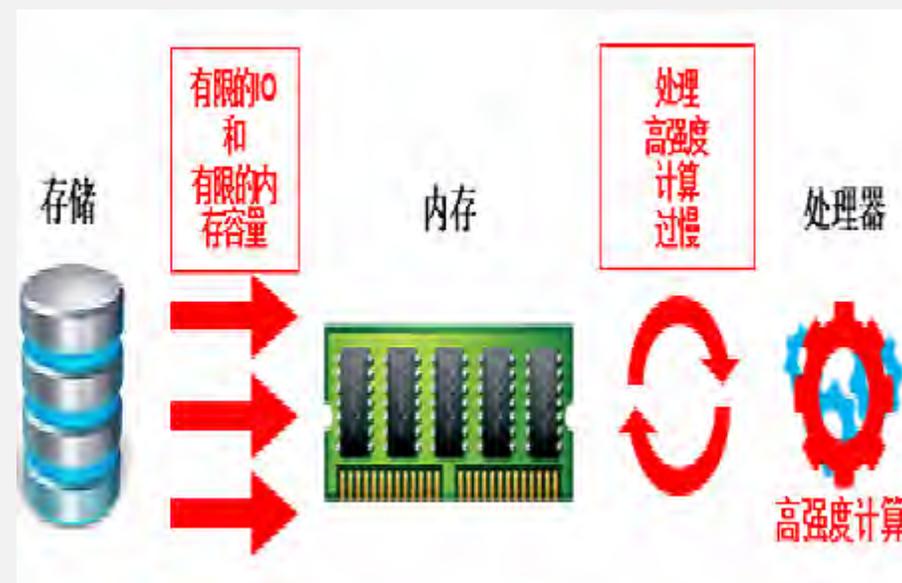


SSD

- 为替换磁盘而有的设计
- 未充分利用闪存的优势
- 无法为应用来定制

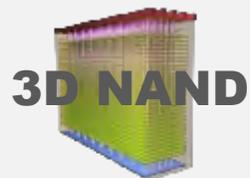
数据中心基础设施的局限性

- 当存储I / O是关键瓶颈的时候，从硬盘驱动器到SSD已经为许多应用提供了有意义的性能提升。
 - SSD在延迟方面实现几个量级的提升，吞吐量亦达到硬盘的10到20倍，对数据中心和云基础设施确实产生了深远的影响。
- 提高存储性能不能解决应用程序对多种密集计算的需求。
 - 对于某一类有明确定义，计算密集型的算法，在现代应用程序的常规数据流中，更适合由其他类型的计算引擎处理。
 - 一些主流算法有例如压缩，纠错码计算和去重，以及其他应用程序相关的功能，如搜索，过滤和位级的数据操作等，能在基于硬件上固化的引擎更有效地执行。
 - 对这种异构计算模型（多个处理数据的引擎架构）的挑战是随着各种应用程序的扩展，特定算法会不断变化。因此，尤为漫长的开发周期使得无法将这些工作负载转移到运行效果更佳的硬件ASIC上。



传统服务器架构对扩展带来的限制

云计算和数据中心基础设施的核心趋势



3D NAND

闪存存储将主宰数据中心
闪存将在2018再提升8倍的容量和密度提升



存储

200MB/s
HDDs



2000MB/s+
SSDs

10x



网络

10Gb/s



100Gb/s+

10x



计算

3 GHz



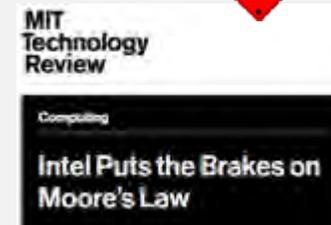
3 GHz



可编程硬件

性能, 成本, 灵活性

2020将有1/3以上的服务器使用FPGA



March, 2016

* Intel Presentation, Dec 2015

性能和能耗推升对FPGA的需求

Amazon EC2 F1 Instances (Preview)

Run Custom FPGAs in the AWS Cloud

F1配有一到八个高端Xilinx现场可编程门阵列（FPGA），为英特尔E5 2686 v4处理器提供可编程硬件，让它可以与高达976G的RAM和4T的NVMe固态硬盘协同工作。FPGA可用于风险管理、模拟、搜索和机器学习应用，或者任何能从经过硬件优化的协同处理器中受益的场景。

Microsoft Accelerates Datacenter With FPGAs

February 23, 2015 by George Leopold

The effort stems from a project launched by Microsoft in 2011 called [Catapult](#) that is intended to leverage FPGAs to **boost datacenter performance, reduce power consumption and develop new datacenter capabilities**...Catapult is Microsoft's response to the anticipated **end of Moore's Law**, the bedrock of current silicon semiconductor design, which a Microsoft researcher predicts is less than one product cycle, "perhaps three years," from hitting a wall...the Catapult platform "**would allow cloud service performance to keep improving, once the silicon scaling hits a wall, by migrating successively larger portions from software into programmable hardware.**"

[云栖社区 > 博客列表 > 正文](#)

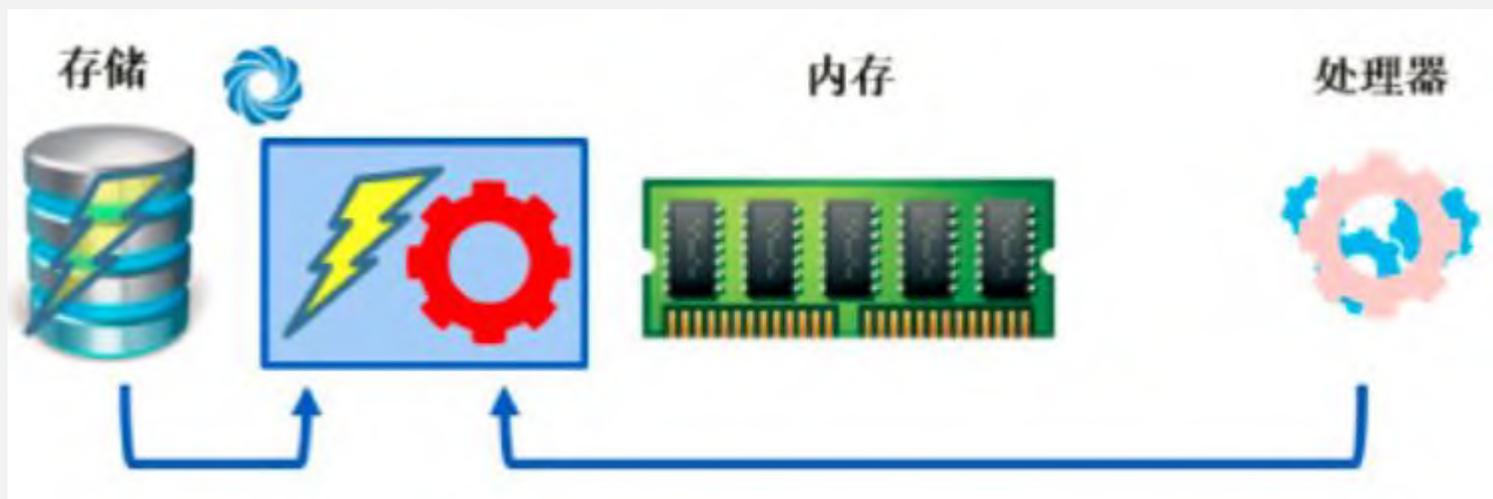
阿里云发布异构计算解决方案：弹性GPU实例及FPGA解决方案

大数据应用暴露出现有数据中心体系上的先天不足

突破性数据流架构

- 计算和存储的融合

- 对密集计算，减少数据从存储器到内存的大量传输。
- 在存储级有效地执行特定的计算功能可以消除大量的数据移动。
- 在所有数据工作负载上用专用计算引擎执行预定义的计算，效率上可以比传统处理器有几个数量级的提高。



闪存和计算引擎的融合

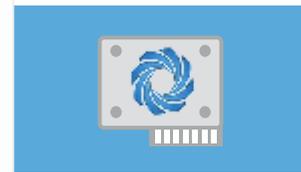
突破性数据流架构

高密度计算
(erasure coding, data scanning)



存储和计算的融合

有限的 I/O 和内存容量



减少数据移动

处理效率提升几十倍

并行计算

重负荷数据移动和计算

应用及分布式文件系统

传统数据计算

- ✓ 大量的数据搬移
- ✓ 重负荷的计算处理
- ✓ 确定性的数据流程



系统资源



融合的云子系统

- 大数据应用的处理
- 缓存功能
- 存储功能
- 网络数据网关功能



软件

特定文件系统和应用程序改变



固件



可编程硬件

整个云存储基础架构的可扩充性

技术壁垒



NAND 闪存 & NVM 技术



Scale Out 横向扩展应用各类算法



领先的 DSP/Error Correction 数据纠错技术



计算和存储系统级架构设计能力



PCIe/NVMe 控制器技术



具突破性的各项专利申请



Thank

You

Jeff Wu

jefewu@outlook.com